1974

The Occurrence of Contrast Effects in Observers' Ratings of Participants in a Leaderless Group Discussion and the Subsequent Reduction of Contrast Effects Through the Provision of Visual and Verbal Standards of Performance

Emmett Lloyd Lipscomb
University of Central Florida

THE OCCURRENCE OF CONTRAST EFFECTS IN OBSERVERS' RATINGS OF PARTICIPANTS IN A LEADERLESS GROUP DISCUSSION AND THE SUBSEQUENT REDUCTION OF CONTRAST EFFECTS THROUGH THE PROVISION OF VISUAL AND VERBAL STANDARDS OF PERFORMANCE

BY

EMMETT LLOYD LIPSCOMB, JR.
B.A., University of South Alabama, 1971

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Psychology
in the Graduate Studies Program of
Florida Technological University

Orlando, Florida
1974

ACKNOWLEDGEMENT

I would like to take this opportunity to express my thanks to the members of my committee, Dr. Edwin Shirkey and Dr. Wayne Burroughs. A special thanks to the Director of this thesis, Dr. Fredric Frank, whose advice and patience made this study possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Originating in Germany by J. B. Rieffert (Ansbacher, 1951), the leaderless group discussion (LGD) has enjoyed widespread use as an assessment technique in the United States since its introduction by the OSS Assessment Staff in the later stages of World War II (Bass, 1954). The practical utility and contribution of the LGD in assessing management potential has been noted by many researchers (e.g., Bass, 1954; Bray and Grant, 1966; Byham, 1969; Jaffee, 1971).

Data on technical efficiency (e.g., interrater reliability, validity, etc.) of the LGD has also been generally favorable. Bass (1954) reports a range of average correlations based on any two raters --using a graphic rating scale--from a minimum correlation of .53 to a maximum of .70. Similarly, a range of correlations from .67 to .90 was reported when checklists, varying from 7 items to 14 items, were used. Estimated reliability correlations, using two raters, ranged from .70 to .82 for graphic rating scales, and .80 to .95 using the checklists. Greenwood and McNamara (1967) report reliabilities from .48 to .82 for ratings, and .48 to .83 for rankings. Bray and Grant (1966), again using two observers, report a correlation of .75 for overall ratings and .75 for overall rankings. However, there is a decrease when observers' ratings are correlated with peer ratings, .69; observers' ratings vs self ratings, .50; and peer vs self ratings, .45. In relation to the interrater reliability, Bray and Grant (1964, p. 10) state "Though relatively high, it is not sufficiently high to warrant dispensing with either of the observers".

While the interrater reliabilities reported above are certainly respectable, interrater reliabilities are not "a priori" indicators

that the raters have accurately judged and recorded the ratees' behavior. In other words, two or more raters may be relatively consistent in their ratings of an individual's behavior and still be consistently in error as to the judged effectiveness or ineffectiveness of that behavior. Far from being an unusual belief, this is an implicit, basic assumption of all research in the area of factors which influence perception and/or judgment. That ratee variables such as physical appearance, sex, race, etc., can and do affect ratings consistently across raters has long been established (e.g., Boehm, 1972; De Jung and Kaplan, 1962; Cox and Krumboltz, 1958). The only thing required is that both raters be susceptible to the biasing variable.

Practically, it would seem that rater bias resulting from variables such as physical appearance, race, and sex could be greatly eliminated through adequate education and training of assessors. However, the degree of precision of a rater's judgment--i.e., how close is the rater's rating to some "absolute" rating--a rating of the individual's performance with all biasing effects excluded--will depend upon the degree to which the rater has clear and precise standards of performance (SP) against which to rate different individuals' behavior, and the degree to which these SP remain constant.

In reference to the SP remaining constant, there are two important aspects. First, the SP being used by several raters should remain constant between raters. If different raters are not using the same SP and are rating the same individual, there will be little comparability or reliability between the ratings. In fact, it seems logical to assume that inconsistent SP between raters is a contributing factor

to less than perfect interrater reliability coefficients, and problems are magnified if different raters having different SP are rating different individuals. In this case, few or no comparisons should be made between the different ratees.

Second, the SP being used by a single rater should remain constant across all individuals he rates. If this interratee or intrarater constancy of SP is not maintained, then there is little validity in trying to compare the performance of individuals who have been rated by the single rater. In effect, different frames of reference may be used for the different ratees, thus ruling out any comparisons.

One might argue that raters within the same organization, for example, should have no problem with uncommon or inconsistent SP because they have received the same training. To some extent this may be true for well trained, professional raters, but the fact is that many companies are training their own employees as raters and these employees are rotated regularly. In other words, the same employees do not always do all the rating. Rather, different groups of employees are used as raters at different times. Also, as Byham (1969) mentions, some companies offer little or no training and others, e.g., AT & T, may have extensive training lasting three weeks. Thus, for those companies offering little or minimum training, a lack of common and consistent SP for raters may lead to errors in ratings.

One such error is the contrast effect. Briefly, a contrast effect is defined as an error in ones' judgment as to the quantity and/or quality of a stimulus due to the changing internal standards of the individual making the judgments, i.e., a discrepancy between the judged

value of the stimulus and its real or "actual" value. This discrepancy may be in a positive or negative direction, e.g., the stimulus may be judged heavier or "better than" it actually is, or it may be judged lighter or "poorer than" it actually is.

The occurrence of contrast effects can be operationally explained through adaptation level (AL) theory (Helson, 1947 and 1964). Basically, AL theory proposes that ones' judgments about the quantity or quality of a stimulus within a series of stimuli reflect that person's adaptation to the whole series of stimuli and not to the single target stimulus. For example, if subjects are given small amounts of training with a target stimulus or training stimulus (S+), and a generalization gradient is then obtained by presenting subjects with a sequence of stimuli and these stimuli are symmetrically distributed about the S+, e.g., a series of weights--5g, 10g, 15g, 20g, 25g, 30g, and 35g, with the 20g weight the S+, maximal responding will be at the S+ value. However, if the same procedure is followed but the stimuli within the sequence are asymmetrically distributed about the S+, e.g., 10g, 15g, 20g, 25g, 30g, 35g, and 40g, with the 20g weight the S+, maximal responding will be found toward the value of the stimuli in the center of the test series. Thomas and Jones (1962) obtained the above results using color stimuli, and Helson and Avant (1967) extended the results to judgments made about the size of pieces of paper. As Guirintano (1973, p. 3-4) states,

> The pooled effects of present and past stimulation establish an internal (subjective) standard against which comparative judgments are made. Rather than a static, isomorphic representation of S+, this internal referent is a dynamic reference point that is sensitive to trial-to-trial changes in stimulation. Consequently, a response (judgment) made on any given trial is relative; it depends upon the relation of stimulation to the momentary AL which has been established...

In other words, an individual's judgment about the quantity or quality of a particular stimulus is dependent upon the context of other stimuli within which it is observed and the resultant frame of reference, AL, established by that context. In effect, within a series of stimuli, each stimulus "tends to pull AL toward its own value, and since the test stimuli comprising an asymmetrical test series are not equally distributed on both sides of the S+, the AL shifts from the S+ toward the center of the test series..." (Guirintano, 1973, p. 3).

The presence of contrast effects in making psychophysical judgments has been verified by many researchers (e.g., Helson, 1947 and 1964; Williams, Ross, and Di Lollo, 1966; Parrish and Smith, 1967) and in the realm of psychosocial judgments, Levine (1972, p. 49) states, "Indeed, many investigators have expressed the conviction that there is a unitary nature, or invariance, of the laws governing psychological judgments". Data supporting this concept of invariance have been reported by quite a few psychosocial studies (e.g., Holmes and Berkowitz, 1961; Rowe, 1967; Hovland, Harvey, and Sherif, 1957; Hakel, Ohnesorge, and Dunnette, 1970; Wexley, Yukl, Kovacs, and Sanders, 1972; and Wexley, Sanders, and Yukl, 1973). So, at this stage of the research there seems little reason to doubt the operation of psychophysical principles in the area of social perception and judgment.

Three of the above studies in the area of psychosocial judgments have particular import for the industrial or personnel psychologist. First is the study by Hakel, et al. (1970). Here, each subject read three employment resumes and was asked to judge the suitability of each of the three applicants. The first two resumes reportedly

provided a frame of reference for the subjects and it was the mean rating of the third resume that was examined for the occurrence of contrast effects. From three qualities of suitability, high (H), average (A), and low (L), six experimental resume sequences were used (HHH, HHL, AAL, AAH, LLH, and LLL). Significant contrast effects were found but they accounted for only 1 to 2% of the total decision variance. It was thus concluded that the contrast effects were only of minor practical significance. However, in the study by Wexley, et al. (1972), results were quite different. In this experiment subjects had a chance to view the supposed applicant in an interview situation. Videotaped interviews were made. Eight sequences of interviews were used (HHH, HHA, HHL, LLH, LLA, LLL, AAH, AAL). Again, the first two videotaped interviews were to establish a frame of reference for the subjects with the mean ratings of the third applicant interview examined for contrast effects. When H or L qualified applicants were viewed in the third position, contrast effects accounted for only a small part of the total variance, i.e., 1 to 2%. But, when applicants of A suitability were rated in the third position, contrast effects accounted for approximately 80% of the total variance. So, in qualifying the Hakel, et al. study, contrast effects were shown to be an important source of practically significant variance in employment interview ratings when average target ratings were evaluated. Third, and most relevant for the present investigation, is a study by Wexley, et al. (1973). Here a series of four experiments was conducted in an effort to eliminate contrast effects in employment interviews. In the first experiment it was found that a warning given to subjects, i.e., "...please make sure that you rate each applicant

on his own merit and not how he compares to those applicants interviewed before him," was ineffective in reducing contrast effects substantially. In the second experiment, anchors describing high and low suitability applicants were supplied to subjects. This was accomplished in the form of written summaries of applicants' responses in an interview situation. Again, contrast effects were not very effectively reduced. In the third experiment a combination of the first two procedures was used. Once again, contrast effects were not substantially reduced. This was somewhat surprising in view of the fact that the two previous procedures were strengthened, e.g., an additional anchor for A suitability was provided to subjects, and they were also given examples of how contrast, stereotyping, leniency, halo, and central tendency could affect ratings. In the fourth experiment a two hour workshop was developed. Subjects were put through this workshop prior to rating the sequence of interviews. Briefly, the workshop "...gave subjects a chance to practice observing and rating actual videotaped applicants, provided subjects with immediate feedback concerning the accuracy of their ratings, and maintained the subjects' interest by using realistic stimuli and by encouraging informal group discussions". This procedure was successful in reducing contrast effects so that they accounted for only 3% of the decision variance. However, due to the methodology used by the experimenters, as acknowledged by the experimenters, no definite conclusions could be drawn regarding which aspects of the workshop were responsible for the great reduction of the contrast effects. In any event, the method used in the workshop would seem to have great import for the training of interviewers.

The occurrence of contrast effects in the interview situation seems to be rather simple and straightforward. However, with reference to LGDs, there seem to be two primary ways for contrast effects to occur. First, contrast effects could occur within a single LGD. For example, how would a participant exhibiting A quality behavior be rated if observed in the context of two or three other participants all exhibiting H or L quality behaviors? Second, contrast effects could possibly occur between LGDs. How would a participant exhibiting A quality performance be rated if the rater had just previously rated two other individuals in two preceding LGDs and their ratings were both H or L? In other words, would a frame of reference or AL be induced in the rater by the first two LGDs and, depending on the rating of the first two individuals, would the frame of reference cause the participant in the third LGD to receive a higher or lower rating than he would have had he not been preceded by the other LGDs?[1]

This study was designed to investigate first, whether contrast effects occur in LGDs as a function of raters having observed previous LGDs. It is hypothesized that, if a target individual participating in an LGD is exhibiting average quality performance and he is preceded by two other target individuals, each participating in other LGDs and both exhibiting and rated as exhibiting high quality performance, then the third target individual will be rated lower than average. Similarly, if the preceding two target individuals are exhibiting and are rated as exhibiting low quality performance, then the third target individual, although actually exhibiting average quality performance, will be rated higher than average. This hypothesis is based on the

data obtained by Wexley, et al. (1972 and 1973) in their investigation of contrast effects in the interview situation.

Second, this study was designed to discover if contrast effects, as defined above, could be reduced or eliminated by supplying standards of performance (SP) to raters. The reduction of contrast effects would be evidenced in two ways. If the third target individual is exhibiting average quality behavior and is preceded by two high target individuals, the third target individual would receive a significantly higher mean rating from raters who had been supplied SP than from raters not having SP supplied. Next, if the third target individual is preceded by two low target individuals, the third target individual should receive significantly lower mean ratings from those raters who had been supplied SP than from the raters not having SP supplied. If contrast effects are completely eliminated by providing SP, the third target individual would receive a mean rating of average regardless of the frame of reference induced by the preceding LGDs.

In reference to the study by Wexley, et al. (1973), the present study differs in three ways. First, the present study was concerned with contrast effects occuring when ratees are performing in an LGD situation. Wexley's, et al.'s study investigated contrast effects when ratees were involved in an interview situation. In the LGD situation presented in this study, the raters actually observed the behaviors which they rated. In the interview situation used by Wexley, et al., the suitability rating given the ratee was not based on actual performance exhibited by the ratee and observed by the raters, but was based on the ratees' responses about his own past performances and

his likes and dislikes. Ten questions asked by the interviewer of the interviewees elicited responses referring to the interviewees' college grade point average, major and minor studies, best and least liked courses, extracurricular activities, leadership positions held, hobbies, past work experience and level of responsibility attained in past jobs, reasons for wanting to be in sales work, feelings about having to travel in the job he was being interviewed for, and what appealed to him most about going to work for the company. So in effect, the performances and experiences on which the interviewees were rated were not exhibited for the raters but were merely reported on.

The second difference between the present study and Wexley's, et al.'s study is the mode of presentation of the anchors or SP. Wexley's anchor stimuli were written summaries of applicants' responses made in interviews. (In Experiment 4 of the Wexley study, subjects did practice and receive feedback on their ratings of individuals in videotaped interviews. However, as previously stated, the experimenters could not directly measure the effect of the anchors. In Experiment 2, in which only written anchors were supplied, the anchors failed to sufficiently reduce the contrast effects. The present study is most similar to Experiment 2 because only anchors were supplied.) The anchor stimuli or SP in the present study were presented visually by means of a videotaped LGD, and verbally by means of the pre-checked or pre-rated rating forms. In the present study then, unlike Wexley's Experiment 2, the mode of presenting the SP to subjects was the same as the mode of presenting those individuals who were to be rated by subjects.

The third difference between this and Wexley's study is that the present study attempts to clearly demonstrate the effects of providing three standards, i.e., high, average, and low. Wexley's Experiment 2 presented only two standards. Experiment 4 provided three standards but did not allow for a direct measure of the effects of the standards.

In summary, this investigation was designed to answer two questions:

1. Do contrast effects, as defined above, occur in the LGD situation provided?
2. Is there a reduction or elimination of contrast effects, as defined above, when SP (H, A, L) are supplied to raters?

Method

Subjects

Subjects were ninety undergraduate students at Florida Technological University. They were recruited from introductory psychology classes and their participation was in partial fulfillment of course requirements. The age range of subjects was from 17 years to 40 years with a mean age of 19.41. There were 47 males and 43 females.

Instruments

A total of nine LGDs were videotaped. Two LGDs contained target individuals exhibiting high (H) quality performance, two contained target individuals exhibiting low (L) quality performance, and three tapes contained target individuals exhibiting average (A) quality performance. A practice videotape was made with undefined performances exhibited by the participants, i.e., there was no single target individual and no particular quality of performance was required of the

participants. Performances exhibited in this practice tape were left undefined for subjects in Group I (see Procedure) so as to decrease the probability of providing them any external SP which they might use. A SP videotape was made with H, A, and L quality performances exhibited by the participants, a single quality for each participant. Each target individual was instructed as to the quality of performance he was to exhibit. See Appendix A for the list of suggested behaviors given to participants in the SP videotape. The suggested behaviors were arrived at by the experimenter's decision. These suggestions were also given to target individuals appearing in other LGDs.

The LGDs consisted of three males and in each LGD the target individual was seated between the other two participants. All but one of the participants were graduate students enrolled in the clinical or industrial psychology programs at the University. The single, non-student participant was a technician employed with the psychology department. Each of the LGDs was twelve to thirteen minutes duration.

Three sequences of LGDs with three LGDs per sequence were videotaped. Within an LGD sequence, different participants were used in each LGD. However, across sequences, the three participants were always the same in the first LGD, in the second LGD, and in the third LGD. A single tape was used for the third LGD in all three sequences. Target individuals also remained the same across sequences, as did their seated position relative to the other two participants in each LGD--seated between the other two participants. Six other participants were used to make the SP videotape and practice LGD--three participants per videotape. Thus, a total of fifteen participants was used in the LGDs.

The task in each LGD was the same. Participants were given a list of ten occupations and instructed to independently rank the occupations from one to ten in order of the prestige they believed was associated with each occupation. After the independent ranking was completed, participants were then instructed to arrive at a final, single ranking by discussing it among themselves. It was the group discussion which was videotaped and shown to subjects. See Appendix B for the instructions, discussion task, and the list of occupations given to the participants.

The rating forms provided to all subjects, with instructions attached, contained fifteen subsections on which the target individuals were to be rated.[2] These subsections were: general activity level, thoroughness, adaptability, willingness to decide, time perspective, commitment, problem analysis, planning and organizing, reaction from others, forcefulness and motivation to lead, style of leading, effectiveness, attitude toward others, oral communication, and reaction to conflict. These scales were constructed from a Goals Checklist provided by Jaffee (1971) for the purpose of evaluating participants in an LGD. However, the scales used were in no way a complete representation of those behaviors presented by Jaffee, but were considered to be the most important areas on which an individual should be rated in an LGD of the short duration used in the present study. The above scales are also in agreement with what line managers in different organizations consider as "...the skills necessary for effectively interacting with other people to get a job done" (Jaffee, 1971). A single, extremes anchored, nine-point scale was used for each

subsection. At the end of the fifteen subsection scales, a single, nine-point overall rating scale was provided. This overall rating scale had extremes and midpoint anchored as excellent, average, and poor. See Appendix C for the rating form and instructions.

A Sony series 3200 camera, Sony series 3650 videotape recorder, Sony series CVT, 19" monitor, and an Altec studio microphone, model 688 B, were used to record and show the LGDs.

Procedure

The general task was to have subjects view a sequence of three LGDs and to rate the quality of performance exhibited by a single target individual in each LGD. Videotaped LGDs were used, rather than written descriptions of behaviors, in order to more closely approximate the actual assessment situation.

Two groups of subjects were used. Both groups viewed a total of four LGDs. The target individual in each of the last three LGDs was rated. Subjects were instructed that the target individual in each LGD was the central participant--the one seated between the other two participants. See Appendix D for general instructions given to both groups. All instructions provided to subjects were read aloud by the experimenter while the subjects read along silently.

LGD sequences and conditions of supplied SP were randomly assigned to groups of subjects as they arrived to participate in the experiment.

For Group 1, no SP were supplied. Subjects in this group were first distributed rating forms and instructions. They were told to observe the practice tape and, during the course of the tape they were to look through one of the rating forms to notice the different

subscales that they would be using to rate target individuals in the following LGDs. The subjects then viewed the practice tape. The practice tape was used to acquaint the subjects with the type of situations they would be viewing, i.e., LGD, the kinds of behaviors they would be rating, and the rating forms they would be using. However, since Group 2 subjects would first be viewing a SP videotape in which standards would be provided for their use in rating target individuals in the following three-LGD sequence, the practice tape for Group 1 was also used as an effort to equate the experience of the two groups with reference to controlling for practice effects which might occur in Group 2 due to their viewing the SP tape. After viewing the practice tape, Group 1 subjects viewed the first LGD and rated the target individual, then viewed the second LGD and rated the target individual, and then viewed the third LGD and rated the target individual. There was an interval of from three to four minutes between LGDs.

Group 2 subjects, as mentioned above, first viewed the SP tape. The purpose of this tape was to provide absolute standards which Group 2 subjects would use in rating the target individuals in the three following LGDs. A standard for H, A, and L quality performance was provided. Immediately following subjects' viewing of the SP tape, the experimenter gave a brief verbal presentation of reasons why each of the participants in the tape should receive their respective ratings of H, A, and L. These reasons consisted of a list of some of the most prominent behaviors (see Appendix E) exhibited by the SP tape participants which was indicative of their respective ratings. Just prior to actually viewing the SP tape, subjects were also provided with

pre-completed rating forms for each of the participants in the SP tape indicating how each of them should have been rated on each of the fifteen subscales and the overall rating. The ratings on the different subscales were determined solely by the experimenter. Verbal agreement on the overall ratings was obtained from several other graduate students that viewed the tapes. See Appendix F for the pre-completed rating forms. Subjects then viewed the sequence of three LGDs and rated the target individual in each LGD at the conclusion of each LGD. None of the subjects in either group were given any training or instructions in combining the different subsections in the rating form.

The target individuals in the LGD sequences exhibited either H, A, or L quality performance. The first two target individuals in each sequence always exhibited the same quality performance, i.e., HH, LL, or AA. The purpose was to induce a frame of reference or AL for subjects. The final or third target individual in each sequence exhibited A quality performance. The same A videotape was used in the third position for all three sequences. Thus, there were three sequence conditions, HHA, LLA, and AAA.

Statistical Analysis

A two-way, fixed effects ANOVA was performed. One independent factor was the frame of reference or AL established in subjects by the target individuals viewed in the first two LGDs in each sequence. A second independent factor was the level of supplied SP. There were two levels--supplied and not supplied. The dependent variable was the subjects' overall rating of the target individual viewed in the third LGD in each sequence.

As there were three sequences of LGDs and two levels of SP provided--no SP provided for Group 1 subjects and three SP supplied for Group 2 subjects--this resulted in the 3 x 2 matrix presented below.

Standards of Performance

LGD Sequences	Not Supplied	Supplied
HHA		
LLA		
AAA		

Results

The mean ratings for target individuals viewed in the first two videotapes in each sequence, and both conditions, standards, and no standards, were 7.23 and 7.13 for H's, 1.79 and 3.16 for L's, and 5.00 and 5.93 for A's. While there is a sizable difference of 1.37 between the mean ratings for the adjacent L's, the data to be presented suggest that this difference was not large enough to inhibit the inducement of an adaptation level (AL) or frame of reference in subjects.

In Table 1 are presented the means and standard deviations of ratings for the target individual in the third LGD in each sequence--the participant exhibiting average (A) quality behaviors--by condition, i.e., standards of performance (SP) provided subjects or no SP provided subjects and the particular sequence in which the A target individual was viewed. Figures 1 and 2 depict graphically the simple main effects of frame of reference (LGD sequence) and SP respectively, on the mean ratings.

TABLE 1

MEAN AND STANDARD DEVIATION OF RATINGS FOR THE TARGET INDIVIDUAL IN THE THIRD LGD IN EACH SEQUENCE

	Sequence	$\bar{X}$	SD
No Standards Provided	HHA	3.13	1.10
	LLA	6.73	1.00
	AAA	4.80	1.10
Standards Provided	HHA	3.80	.812
	LLA	5.33	1.50
	AAA	4.46	1.05

Note: Abbreviations--H = high quality behavior exhibited by target individual, A = average quality behavior exhibited by target individual, L = low quality behavior exhibited by target individual.

FIGURE 1

PROFILES OF SIMPLE MAIN EFFECTS FOR SEQUENCES

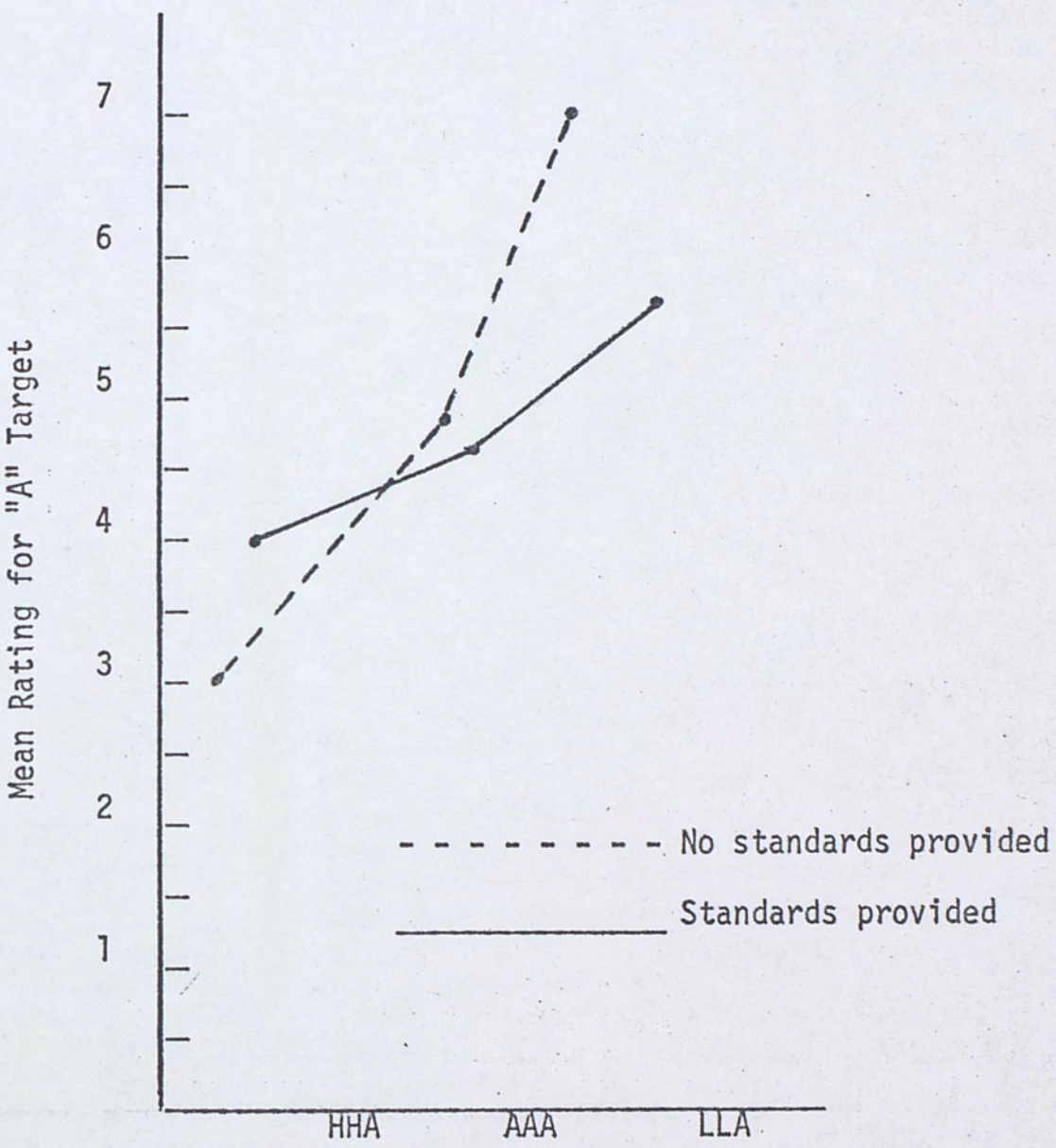
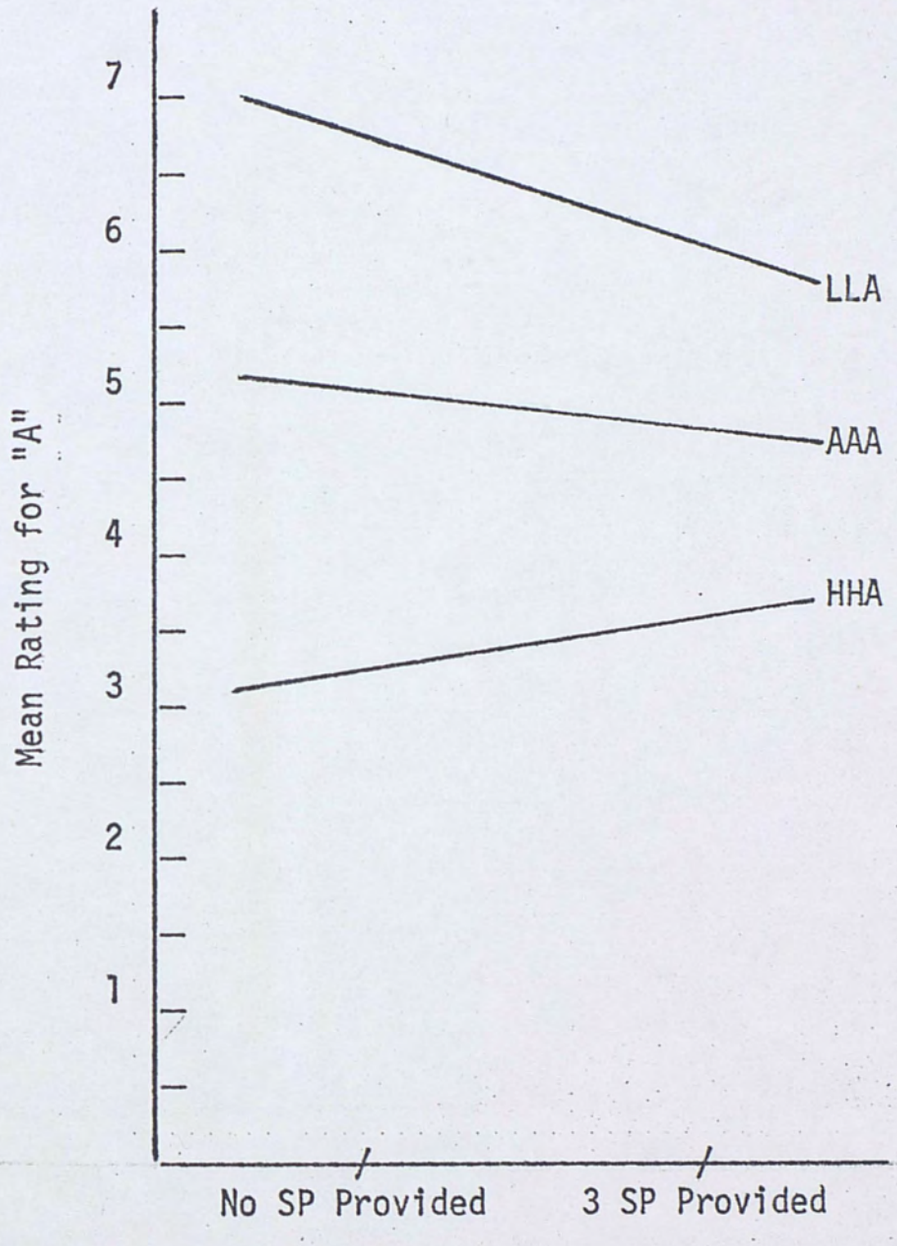


FIGURE 2

PROFILES FOR SIMPLE MAIN EFFECTS FOR SP



Interaction between frame of reference (H, A, and L) and SP provided (0 or 3) was examined initially as this was of primary importance. First, differences between sequences, in the ratings of the third target individual, when no SP were provided would demonstrate the occurrence of contrast effects. Second, if the differences just mentioned were not present when three SP were provided, reduction or elimination of contrast effects would be demonstrated. The ANOVA results as summarized in Table 2 do indicate a significant interaction effect ($F = 5.80$, $p < .01$, 2 7 84 df). The presence of significant interaction indicates that the effect of frame of reference induced in raters on ratings is dependent upon provision of SP (providing no SP or providing three--H, A, and L). In other words, the effect that induced frame of reference had on a subject's rating of the target individual in the third LGD was dependent on whether or not that subject had been provided SP on which to base his rating.

A computation of simple effects (Winer, 1962) was conducted to determine the nature of the interaction. Table 3 summarizes the analysis to determine the simple effects of sequence (frame of reference) on SP provided--none or three. It is indicated that there were significant simple effects of frame of reference when no SP were supplied ($F = 35.08$, $p < .01$, 2 & 84 df), and also when the SP were supplied ($F = 6.42$, $p < .01$, 2 & 84 df).

Post hoc comparisons were made to determine the relationship between SP supplied and the rating of the third target individual in the HHA, LLA, and AAA sequences, and between no SP supplied and the rating of the third target individual in the HHA, LLA, and AAA sequences. As

TABLE 2

ANALYSIS OF VARIANCE OF RATINGS FOR THE TARGET INDIVIDUAL IN THE THIRD LGD IN EACH SEQUENCE

Source	SS	df	MS	F	W^2
Row (Frame of Reference	99.09	2	49.54	35.89**	.408
Column (SP)	2.84	1	2.84	2.05	.006
Interaction	16.02	2	8.01	5.80 *	.056
Error (Within Cell)	116.54	84	1.38		
Totals	234.49	89			

* $p \leq .01$

** $p \leq .001$

TABLE 3

ANALYSIS OF VARIANCE OF THE SIMPLE EFFECTS OF FRAME OF REFERENCE ON STANDARDS OF PERFORMANCE

Source	SS	df	MS	F	W^2
FR for no SP	97.38	2	48.69	35.28 *	.403
FR for 3 SP	17.73	2	8.86	6.42 *	.063
Error (Within Cell)	116.54	84	1.38		

* $p \leq .01$

Note: Abbreviations--FR = frame of reference, SP = standards of performance.

indicated in Table 4, when no SP were supplied contrast was in evidence. The third rating in the HHA sequence was significantly lower than the third rating in the LLA sequence ($F = 54$, $p < .01$, 84 df) and significantly lower than the third rating in the AAA sequence ($F = 25$, $p < .01$, 84 df). Also, the third rating in the LLA sequence was significantly higher than the third rating in the AAA sequence ($F = 29$, $p < .01$, 84 df).

The fact that frame of reference also had significant simple effects when SP were supplied to subjects indicates the lack of complete effectiveness of the technique used to supply the SP. The reasoning is that if providing SP was completely effective, then there should be no simple effect of frame of reference when three SP are provided. As indicated in Table 5, when SP were supplied to subjects, the third rating in the LLA sequence was significantly higher than the third rating in the AAA ($F = 13$, $p < .05$, 84 df) and HHA sequences ($F = 23$, $p < .01$, 84 df). The third rating in the HHA and AAA sequences were not significantly different.

Table 6 summarizes the analysis to determine the simple effects of SP on sequence or frame of reference. As indicated in Table 6, SP had a significant simple effect only on the LLA sequence, i.e., the L frame of reference provided subjects ($F = 10.65$, $p < .01$, 1 & 84 df). By referring to Table 1, one sees that the effect was to significantly reduce the mean rating of the third target individual from 6.73 when no SP were provided, to 5.33 when three SP were provided. While the change in mean ratings for the HHA sequence (H frame of reference) was

TABLE 4

CELL TOTAL DIFFERENCES BETWEEN SEQUENCES

WITH NO STANDARDS OF PERFORMANCE SUPPLIED

	HHA	AAA	LLA
HHA	---	25 *	54 *
AAA		---	29 *
LLA			---

* $p < .01$

Critical value for $r = 1$, 17.13

$r = 2$, 19.50

TABLE 5

CELL TOTAL DIFFERENCES BETWEEN SEQUENCES WITH HIGH, AVERAGE, AND LOW STANDARDS OF PERFORMANCE SUPPLIED

	HHA	AAA	LLA
HHA	---	10	23**
AAA		---	13 *
LLA			---

*p < .05

**p < .01

Critical value for .05: r = 1, 12.89

r = 2, 15.49

Critical value for .01: r = 1, 17.13

r = 2, 19.50

TABLE 6

ANALYSIS OF VARIANCE OF THE SIMPLE EFFECTS OF STANDARDS OF PERFORMANCE ON FRAME OF REFERENCE

Source	SS	df	MS	F
SP for FR-H	3.33	1	3.33	2.41
SP for FR-L	14.70	1	14.70	10.65*
SP for FR-A	.83	1	.83	.60
Error (Within Cell)	116.54	84	1.38	

*$p \leq .01$

Note: Abbreviations--FR-H = high frame of reference provided, FR-L = low frame of reference provided, FR-A = average frame of reference provided.

not significant, it was in the desired direction, i.e., when SP were provided the mean rating was increased from 3.13 to 3.80--closer to a rating of "Average". In the third sequence (AAA), as expected, the decrease in the mean rating when SP were supplied was not significant. In effect, if providing SP was to be effective overall, there should have been significant simple effects of SP on the HHA sequence as well as the LLA sequence.

Omega squared was computed to determine the amount of variance accounted for by sequence, i.e., frame of reference. Table 3 indicates that when the ANOVA of the simple effects of frame of reference on standards of performance is considered, 40.3% of the total variance was due to the condition of no standards provided to subjects. When three SP were provided, frame of reference accounted for only 6.3% of the variance. So, while some significant differences were present when SP were provided, they accounted for only 6.3% of the total variance.

To summarize, contrast effects were produced in the LGD's and the contrast effects were substantially reduced but not totally eliminated by providing standards of performance.

Discussion

Two findings are of considerable interest. First, contrast effects can be and are, in the situation examined in this investigation, a statistically significant source of variance in ratings of individuals who exhibit average quality behaviors while participating in LGDs.

The second characteristic of contrast effects as indicated by this study's data, and supported by that study of Wexley, et al. (1973), is that while contrast effects can be a significant source of variance in ratings, the effects can be reduced.

The amount of total variance due to the contrast effects was reduced to less than 7% by providing standards, compared to 40% when standards were not provided. However, it was noted that contrast was not significantly reduced in the HHA condition. One possibility is that the SP presented subjects via the SP film were not equally explicit, i.e., the SP did not supply equally effective points of reference along the excellent--average--poor continuum. It may be noted in Appendix E that more behaviors were listed to describe the "excellent" behavior than the "average" or "poor" behavior. This may have resulted in a "stronger" standard for "excellent" than for "poor" or "average". If it is actually true that the "excellent" standard was more effective than the "poor" or "average" standard, then, hypothetically, when subjects viewed the A target individual in the LLA sequence the tendency to rate the target higher than average was inhibited due to the explicitly defined standard for "excellent" behavior. When subjects viewed the HHA sequence the tendency to rate the A target individual lower than average was not reduced to the same degree, as that tendency produced in the LLA sequence, because the poor standard was not as fully defined as the "excellent" standard. In other words, the subjects were more sure of what was not "excellent" than of what was not "poor".

Another possible explanation is based on a possible difficulty on the subjects' part in trying to combine the different subscales on the

pre-checked rating forms. It is indicated in Appendix F that there was considerable fluctuation in the numerical ratings on the subscales for the "average" target individual. Including a "does not apply" (DNA) response, there were seven different numerical ratings ranging from "three" to "nine". There were five different ratings ranging from "one" to "five" for the "poor" target and only two different ratings, "eight" and "nine", for the "excellent" target. This would seem to relate back to the explicitness of the standards. It would seem reasonable to propose that a rater would find it easier to arrive at an overall rating for an individual that is consistently exhibiting "excellent" quality behaviors--evidenced by a range of only two scale points--than for an individual whose exhibited behavior is inconsistent and fluctuates over a range of five scale ratings--as did the "poor" standard target individual. The "average" standard fluctuated over six different scale ratings. Presumably, the more fluctuation in the exhibited behavior of the standard, the less precise or explicit is the definition of that standard. The fact that the subjects were not given any instructions on how to combine the subscales to arrive at an overall rating may have affected their ability to combine the scales. However, this does not seem to have been an artificial effect because there is generally a large degree of fluctuation in the performance of most individuals--especially those individuals considered to be "average".

While two possible explanations have been offered for the small reduction of contrast effects in the HHA sequence--compared to the LLA sequence--it should be reiterated that, when SP were not provided, significant contrast effects were found in both the LLA sequence and

the HHA sequence. More importantly, overall, contrast was of minimal impact when SP were provided. Clearly while the SP were not totally effective, they were more effective than the Warning (Experiment 1), the Anchoring (Experiment 2), or the Combination (Experiment 3), as investigated by Wexley, et al. (1973). They were not as effective as the Workshop (Experiment 4) in which contrast effects were reduced to accounting for only 3% of the variance.

The results of this study have considerable practical significance. To maximize the efficient utilization of all employees, or prospective employees, it is necessary that each employee be judged against common, consistent standards. The results of this present investigation, along with those by Wexley, et al. (1972 and 1973), bring into question the degree to which the use of common, stable standards has been accomplished. Rather, it has been demonstrated that it is quite possible for standards to fluctuate and thus affect the accuracy of ratings. As mentioned above, when no definite standards were provided the raters, the amount of variance in the ratings due to contrast was approximately 40%. In an ideal situation, all of the variance should be due to the actual quality of the behaviors exhibited regardless of the sequence in which those behaviors may be observed.

The results of this study are particularly relevant to those involved in assessment through the use of the LGD. While it has been shown that contrast effects can present a problem when rating individuals participating in an LGD, it has also been shown that at least the problem can be reduced to minimal proportions.

Too direct an application of this investigation's results may be premature. First, the mean age of the subjects was only 19.4 years. Subjects of this age generally have little or no experience in the formal rating of others--especially in an LGD. It does not seem likely that nineteen year old persons would be used as raters in the actual industrial situation. Further, while some managers or supervisors may not be familiar with the formal LGD, practically all would have had some experience in the formal appraisal systems for subordinates. However, the use of subjects dissimilar to the population of actual concern has become a practical necessity for the great majority of experimenters.

Second, in this investigation the time span between subsequent LGDs was only three to four minutes. In the industrial assessment situation there is often a much longer break between LGDs, e.g., one day or more. This could allow for a sufficient restructuring of standards of performance within the raters. More clearly stated, since there may be a day or more between the different LGDs, it may be that a fixed or permanent frame of reference or adaptation level is not induced. However, the increased time interval does not preclude the establishment of a frame of reference while the rater is viewing a single LGD and making comparisons among the participants in that LGD. Further research involving the manipulation of the time interval between the viewing of LGDs is needed to determine two things. First, do contrast effects occur within a single LGD; second, if contrast effects do occur in a single LGD, then, as the time interval between successively viewed LGDs increases, do contrast effects within each LGD account for more of the decision variance than contrast effects between adjacent LGDs.

With further refinement it is thought that contrast effects could be entirely eliminated. At the present time, however, further research utilizing visual and verbal SP is necessary. Of particular interest for practical application is the determination of the effectiveness of visual and verbal SP over a period of time. Does the effectiveness of presenting definite SP decay over time, i.e., is it necessary to present the SP at short intervals between the viewing of a particular number of LGDs in order to maintain the effectiveness of reducing or eliminating contrast effects? Of course, of prime interest is the eventual research which must be done in the actual industrial situation.

APPENDICES

APPENDIX A

SUGGESTED BEHAVIORS FOR PARTICIPANTS IN SP VIDEOTAPE

APPENDIX A

QUALITY OF PERFORMANCE TO EXHIBIT - EXCELLENT

1. Request that the group have a chairman.
2. Initiate the group activity - poll the group to determine how they stand.
3. Assume control of the group - try to run the group - be directive but not authoritarian - make some decisions.
4. Be polite and congenial.
5. Speak clearly and distinctly.
6. If disagreed with maintain your composure. If disagreed with defend your position vigorously.
7. Address others by name.
8. Take some notes during all part of the discussion.
9. At some point later in the discussion take notice of how much time is left to complete the problem.
10. Try to present your views in a well organized manner.
11. Press for a final group decision before time runs out.
12. Make a motion toward the end of the exercise that will end the discussion.
13. Be as involved, as participative, as possible - be energetic.
14. If your proposal is related to another's proposal, mention it.
15. If it is difficult to change someone's mind, note the difficulty.

APPENDIX A (Con't.)

QUALITY OF PERFORMANCE TO EXHIBIT - AVERAGE

1. Don't request that the group have a chairman, but agree with the request when it is made by another.
2. Don't initiate any group activity but follow the lead of the one who does initiate the activity.
3. Try to gain control of the group but fail - give in to the leader.
4. Address others by name - look at others when speaking to them.
5. Be polite, congenial.
6. Speak clearly, distinctly.
7. If disagreed with become a little irritated, but not angry. If disagreed with, defend your position moderately and give in to what the leader proposes.
8. Take notes during the first part of the discussion but then quit taking notes.
9. At no time during the discussion take any notice of how much time is left to complete the problem.
10. Try to present your views in an orderly manner.
11. Don't press for a final group decision before time runs out but agree with and follow the leader of the group when he does so.
12. Don't make a motion that would end the exercise but again, agree with and follow the leader.
13. Be moderately participative - keep involved in the discussion.

APPENDIX A (Con't.)

QUALITY OF PERFORMANCE TO EXHIBIT - POOR

1. Don't request that the group have a chairman.
2. Don't initiate any activity.
3. Don't try to gain control of the group - be passive, don't make any decisions.
4. Don't address others by name - don't look at others when you are speaking to them.
5. Be a little rude, sarcastic, unfriendly.
6. Mumble your words - be indistinct.
7. If disagreed with become angry, lose your composure. If disagreed with don't defend your position.
8. Don't take any notes during the discussion.
9. Don't consider how much time you have left - make no notice of it.
10. Present your views in an organized, rambling manner.
11. Don't press for a final group decision.
12. Don't make a motion toward the end of the exercise that will end the discussion.
13. Be as non-participative as possible. Never really become involved in the discussion.

APPENDIX B

INSTRUCTIONS AND TASK

GIVEN TO LGD PARTICIPANTS

APPENDIX B

INSTRUCTIONS

I. Below is a list of ten occupations. Your task is to rank these occupations, by yourself, in order of the prestige you think is associated with each.

Are there any questions?

Physician

Lawyer

College Professor

Engineer

Computer Specialist

Nuclear Physicist

Psychiatrist

Psychologist

Certified Public Accountant

Marine Biologist

II. It is now your task to arrive at a single, final ranking of the occupations. You are to do this by discussing what you have with the other group members. You have approximately twelve minutes to complete this task.

Are there any questions?

APPENDIX C

RATING FORM,

RATING FORM INSTRUCTIONS

AND ADDENDUM

APPENDIX C

LGD RATING FORM

A. INDIVIDUAL WORK CHARACTERISTICS

1. General Activity Level

9 8 7 6 5 4 3 2 1

Always involved in the discussion — Never involved in the discussion

2. Thoroughness

9 8 7 6 5 4 3 2 1

Was thorough in working on the problem — Was not thorough in working on the problem

3. Adaptability

9 8 7 6 5 4 3 2 1

Was flexible in his problem solving approach — Was not flexible in his problem solving approach

B. DECISION MAKING STYLE

1. Willingness to Decide

9 8 7 6 5 4 3 2 1

Was willing to make decisions — Was not willing to make decisions

2. Time Perspective

9 8 7 6 5 4 3 2 1

Was aware of the importance of the time element — Was not aware of the importance of the time element

APPENDIX C (Con't.)

3. Commitment

9 8 7 6 5 4 3 2 1

Was willing to defend his positions — Was not willing to defend his positions

C. ORGANIZATION AND PLANNING STYLE

1. Problem Analysis

9 8 7 6 5 4 3 2 1

Was aware of the difficulties involved — Was not aware of the difficulties involved

2. Planning and Organizing

9 8 7 6 5 4 3 2 1

Was organized and planned ahead — Was not organized and did not plan ahead

D. LEADERSHIP BEHAVIOR

1. Reaction from Others

9 8 7 6 5 4 3 2 1

Was responded to well by the others — Was responded to poorly by the others

2. Forcefulness and Motivation to Lead

9 8 7 6 5 4 3 2 1

Was determined to lead the group - aggressive — Was not determined to lead the group - passive

APPENDIX C (Con't.)

3. Style of Leading

9 8 7 6 5 4 3 2 1

Was democratic in trying to direct the group

Was authoritarian in trying to direct the group

4. Effectiveness

9 8 7 6 5 4 3 2 1

Group usually followed his directions and opinions

Group usually did not follow his directions and opinions

E. INTERPERSONAL CHARACTERISTICS

1. Attitude Toward Others

9 8 7 6 5 4 3 2 1

Was cooperative, polite to others in the group

Was uncooperative, impolite to others in the group

2. Oral Communication

9 8 7 6 5 4 3 2 1

Was easily understood

Was not easily understood

3. Reaction to Conflict

9 8 7 6 5 4 3 2 1

Did not become upset when disagreed with by others

Became upset when disagreed with by others

APPENDIX C (Con't.)

OVERALL RATING

How would you rate the overall quality of this individual's performance in this exercise?

9	8	7	6	5	4	3	2	1
Excellent				Average				Poor

APPENDIX C (Con't.)

RATING FORM INSTRUCTIONS

The following Rating Form is divided into subsections. A single scale is provided for each subsection - a total of fifteen scales. From what you have observed, rate the quality of the performance exhibited by the individual you have observed.

Each scale goes from 1 to 9. Circle the number you think best represents the quality of performance of the observed individual. The higher the quality of performance, the higher should the circled number be. The lower the quality of performance, the lower should the circled number be.

After you have completed your rating on the fifteen subscales, you will find an <u>overall</u> rating scale. Circle the number which you think best represents the <u>overall quality</u> of the observed individual's performance.

Are there any questions?

APPENDIX C (Con't.)

ADDENDUM TO APPENDIX C

It was determined by the experimenter that all of the fifteen subscales would not apply to all of the target individuals. For example, if a target individual was exhibiting "poor" quality performance, it is unlikely that subjects would have any way to rate the target on either the adaptability or style of leading subscales. Subjects could not rate the individual on adaptability (note anchors for extremes on this subscale) because the individual may have had no problem solving approach. Similarly, for the style of leading subscale, if the target individual at no time tried to lead the group, then subjects could not rate him on whether he was a democratic or authoritarian leader.

For the above reasons the subjects were presented the following verbal instructions:

"You are to use as many of the subscales as possible when rating the LGD participants. However, if for some reason you feel that a particular subscale does not apply to a particular participant, please print "DNA" to the left of that particular scale. Again, try to use as many of the subscales as possible.

Are there any questions?"

APPENDIX D

GENERAL INSTRUCTIONS, GROUPS I AND II, RESPECTIVELY

APPENDIX D

GENERAL INSTRUCTIONS

Part 1:

During this experiment you will view four videotaped films. The first film you will see is simply a practice film to acquaint you with the kind of exercise you will be viewing and the kinds of behavior you will be observing.

Three persons will be participating in each of the films you will view. The participants are involved in a leaderless group discussion (LGD). Before these films were made, each of the three participants was given a list of ten occupations. It was their task to individually rank these occupations from 1 to 10, in reference to the prestige associated with the occupations. After this was done on an individual basis, the participants were instructed to arrive at a single, final ranking of the occupations by discussing it among themselves. It is this discussion that you will be observing in each of the films.

After this practice film is over, look at the rating form you have been provided and notice the different scales you will be using to make your ratings in the following films. Instructions are attached to the rating form.

Are there any questions?

Part 2:

In the next three films, concentrate your attention on the person seated in the central position--seated between the other two participants. Immediately after each discussion is concluded, rate the quality of the

APPENDIX D (Con't.)

central individual's performance on the rating form provided. Instructions for use of the rating form are attached to the form.

Are there any questions?

APPENDIX D (Con't.)

GENERAL INSTRUCTIONS

Part 1:

During this experiment you will view four videotaped films. The first film you will view is a "Standards of Performance" film. This film is designed to provide you with standards which you can use in rating particular individuals in three following films.

Each individual in the "Standards" film will be exhibiting a different quality of performance. Individual A, seated on your left, will exhibit excellent quality performance. Individual B, seated in the center position, will exhibit average quality performance. Individual C, seated on your right, will exhibit poor quality performance.

When the film is over, you are to look at the checked rating forms that have been provided for each of the individuals; i.e., A, B, and C. These rating forms indicate the correct rating for each of the individuals on each of the scales. You are to use the ratings on these forms and the corresponding behaviors you observed in the film as standards of performance when you rate the following individuals in the next three films.

In all of the films you will be viewing, the three participants in each film will be involved in a leaderless group discussion (LGD). Before these films were made, each of the three participants was given a list of ten occupations. It was their task to individually rank these occupations from 1 to 10, in reference to the prestige associated with the occupations. After this was done on an individual basis, the

APPENDIX D (Con't.)

participants were instructed to arrive at a single, final ranking of the occupations by discussing it among themselves. It is this discussion that you will be observing in each of the films.

Are there any questions?

Part 2:

In the next three films, concentrate your attention on the person seated in the central position--seated between the other two participants. Immediately after each discussion is concluded, rate the quality of the central individual's performance on the rating form provided. Instructions for use of the rating form are attached to the form.

Remember to rate these individuals by comparing their behaviors to the behaviors you observed in the "Standards of Performance" film and the corresponding ratings which were shown to you. Compare the central individual's behavior to what you saw in the "Standards" film.

Are there any questions?

APPENDIX E

BEHAVIORS EXHIBITED BY

PARTICIPANTS IN STANDARDS OF PERFORMANCE FILM

APPENDIX E

BEHAVIORS EXHIBITED BY PARTICIPANTS IN STANDARDS OF PERFORMANCE FILM

INDIVIDUAL A - EXCELLENT

1. Initiated group discussion.
2. Assumed leadership of the group.
3. Polled other two participants as to their rankings.
4. Supplied organization for completing the task.
5. Was involved in the discussion practically all of the time.
6. Addressed others by name.
7. Voiced his opinions and defended them.
8. Was democratic in his style of leading - asked for and listened to the opinions of the other participants.
9. Got other participants to agree to his decisions most of the time.
10. Noticed that the group was running out of time and pressed for completion of the task.
11. Did not let disagreements bother him - maintained his composure.
12. Spoke clearly and distinctly.
13. Took notes throughout the discussion.

APPENDIX E (Con't.)

INDIVIDUAL B - AVERAGE

1. Did not initiate discussion but was involved in the discussion most of the time.
2. Did not try to gain control of the group.
3. Voiced his opinions and defended them.
4. Got Individual A to accept his position at least once.
5. Did not become upset when he was disagreed with.
6. Was flexible - did not demand that his positions be accepted.
7. Took notes throughout most of the discussion.
8. Spoke clearly and distinctly.
9. Did not take notice that the group was running out of time but did follow Individual A's press for completion of the task before time ran out.

APPENDIX E (Con't.)

INDIVIDUAL C - POOR

1. Never initiated discussion.
2. Never became involved in the discussion - did not really try to work on the problem.
3. Did not try to gain control of the group.
4. Did not try to make any decisions.
5. Went along with whatever positions the other participants wanted - very passive.
6. Did not keep up with what the group was working on - got lost as to what occupations were being discussed.
7. Spoke in low voice - rather indistinct.
8. Doodled and gazed about the room while the discussion was in progress - seemed very uninterested in the discussion.
9. Did not notice how much time was left and tried to continue the discussion when Individual A pressed for completion of the task before time ran out.

APPENDIX F

PRE-COMPLETED RATING FORMS FOR PARTICIPANTS IN THE STANDARDS OF PERFORMANCE VIDEOTAPE

APPENDIX F

LGD RATING FORM - INDIVIDUAL A

A. INDIVIDUAL WORK CHARACTERISTICS

1. General Activity Level

(9) 8 7 6 5 4 3 2 1

Always involved in the discussion — Never involved in the discussion

2. Thoroughness

9 (8) 7 6 5 4 3 2 1

Was thorough in working on the problem — Was not thorough in working on the problem

3. Adaptability

(9) 8 7 6 5 4 3 2 1

Was flexible in his problem solving approach — Was not flexible in his problem solving approach

B. DECISION MAKING STYLE

1. Willingness to Decide

(9) 8 7 6 5 4 3 2 1

Was willing to make decisions — Was not willing to make decisions

2. Time Perspective

(9) 8 7 6 5 4 3 2 1

Was aware of the importance of the time element — Was not aware of the importance of the time element

APPENDIX F (Con't.)

3. Commitment

(9) 8 7 6 5 4 3 2 1

Was willing to defend his positions — Was not willing to defend his positions

C. ORGANIZATION AND PLANNING STYLE

1. Problem Analysis

9 (8) 7 6 5 4 3 2 1

Was aware of the difficulties involved — Was not aware of the difficulties involved

2. Planning and Organizing

9 (8) 7 6 5 4 3 2 1

Was organized and planned ahead — Was not organized and did not plan ahead

D. LEADERSHIP BEHAVIOR

1. Reaction from Others

9 (8) 7 6 5 4 3 2 1

Was responded to well by the others — Was responded to poorly by the others

2. Forcefulness and Motivation to Lead

(9) 8 7 6 5 4 3 2 1

Was determined to lead the group - aggressive — Was not determined to lead the group - passive

APPENDIX F (Con't.)

3. Style of Leading

(9) 8 7 6 5 4 3 2 1

Was democratic in trying to direct the group

Was authoritarian in trying to direct the group

4. Effectiveness

(9) 8 7 6 5 4 3 2 1

Group usually followed his directions and opinions

Group usually did not follow his directions and opinions

E. INTERPERSONAL CHARACTERISTICS

1. Attitude Toward Others

9 (8) 7 6 5 4 3 2 1

Was cooperative, polite to others in the group

Was uncooperative, impolite to others in the group

2. Oral Communication

(9) 8 7 6 5 4 3 2 1

Was easily understood

Was not easily understood

3. Reaction to Conflict

(9) 8 7 6 5 4 3 2 1

Did not become upset when disagreed with by others

Became upset when disagreed with by others

APPENDIX F (Con't.)

OVERALL RATING

How would you rate the overall quality of this individual's performance in this exercise?

(9) 8 7 6 5 4 3 2 1

Excellent Average Poor

APPENDIX F (Con't.)

LGD RATING FORM - INDIVIDUAL B

A. INDIVIDUAL WORK CHARACTERISTICS

1. General Activity Level

9 8 7 (6) 5 4 3 2 1

Always involved in the discussion — Never involved in the discussion

2. Thoroughness

9 8 7 6 5 (4) 3 2 1

Was thorough in working on the problem — Was not thorough in working on the problem

3. Adaptability

9 8 (7) 6 5 4 3 2 1

Was flexible in his problem solving approach — Was not flexible in his problem solving approach

B. DECISION MAKING STYLE

1. Willingness to Decide

9 8 7 6 (5) 4 3 2 1

Was willing to make decisions — Was not willing to make decisions

2. Time Perspective

9 8 7 6 (5) 4 3 2 1

Was aware of the importance of the time element — Was not aware of the importance of the time element

APPENDIX F (Con't)

3. Commitment

9 8 7 6 (5) 4 3 2 1

Was willing to defend his positions

Was not willing to defend his positions

C. ORGANIZATION AND PLANNING STYLE

1. Problem Analysis

9 8 7 6 5 (4) 3 2 1

Was aware of the difficulties involved

Was not aware of the difficulties involved

2. Planning and Organizing

9 8 7 6 (5) 4 3 2 1

Was organized and planned ahead

Was not organized and did not plan ahead

D. LEADERSHIP BEHAVIOR

1. Reaction from Others

9 8 7 (6) 5 4 3 2 1

Was responded to well by the others

Was responded to poorly by the others

2. Forcefulness and Motivation to Lead

9 8 7 6 5 4 (3) 2 1

Was determined to lead the group - aggressive

Was not determined to lead the group - passive

APPENDIX F (Con't)

3. Style of Leading

DNA 9 8 7 6 5 4 3 2 1

Was democratic in trying to direct the group

Was authoritarian in trying to direct the group

4. Effectiveness

9 8 7 6 (5) 4 3 2 1

Group usually followed his directions and opinions

Group usually did not follow his directions and opinions

E. INTERPERSONAL CHARACTERISTICS

1. Attitude Toward Others

9 8 (7) 6 5 4 3 2 1

Was cooperative, polite to others in the group

Was uncooperative, impolite to others in the group

2. Oral Communication

9 8 (7) 6 5 4 3 2 1

Was easily understood

Was not easily understood

3. Reaction to Conflict

(9) 8 7 6 5 4 3 2 1

Did not become upset when disagreed with by others

Became upset when disagreed with by others

APPENDIX F (Con't.)

OVERALL RATING

How would you rate the overall quality of this individual's performance in this exercise?

9 8 7 6 (5) 4 3 2 1

Excellent Average Poor

APPENDIX F (Con't.)

LGD RATING FORM - INDIVIDUAL C

A. INDIVIDUAL WORK CHARACTERISTICS

1. General Activity Level

9 8 7 6 5 4 3 2 (1)

Always involved in the discussion — Never involved in the discussion

2. Thoroughness

9 8 7 6 5 4 3 2 (1)

Was thorough in working on the problem — Was not thorough in working on the problem

3. Adaptability

DNA 9 8 7 6 5 4 3 2 1

Was flexible in his problem solving approach — Was not flexible in his problem solving approach

B. DECISION MAKING STYLE

1. Willingness to Decide

9 8 7 6 5 4 3 2 (1)

Was willing to make decisions — Was not willing to make decisions

2. Time Perspective

9 8 7 6 5 4 3 2 (1)

Was aware of the importance of the time element — Was not aware of the importance of the time element

APPENDIX F (Con't.)

3. Commitment

9 8 7 6 5 4 3 2 (1)

Was willing to defend his positions — Was not willing to defend his positions

C. ORGANIZATION AND PLANNING STYLE

1. Problem Analysis

9 8 7 6 5 4 3 2 (1)

Was aware of the difficulties involved — Was not aware of the difficulties involved

2. Planning and Organizing

9 8 7 6 5 4 3 2 (1)

Was organized and planned ahead — Was not organized and did not plan ahead

D. LEADERSHIP BEHAVIOR

1. Reaction from Others

9 8 7 6 5 4 (3) 2 1

Was responded to well by the others — Was responded to poorly by the others

2. Forcefulness and Motivation to Lead

9 8 7 6 5 4 3 2 (1)

Was determined to lead the group - aggressive — Was not determined to lead the group - passive

APPENDIX F (Con't.)

3. Style of Leading

DNA 9 8 7 6 5 4 3 2 1

Was democratic in trying to direct the group

Was authoritarian in trying to direct the group

4. Effectiveness

9 8 7 6 5 4 3 2 (1)

Group usually followed his directions and opinions

Group usually did not follow his directions and opinions

E. INTERPERSONAL CHARACTERISTICS

1. Attitude Toward Others

9 8 7 6 5 4 3 (2) 1

Was cooperative, polite to others in the group

Was uncooperative, impolite to others in the group

2. Oral Communication

9 8 7 6 5 4 3 (2) 1

Was easily understood

Was not easily understood

3. Reaction to Conflict

9 8 7 6 (5) 4 3 2 1

Did not become upset when disagreed with by others

Became upset when disagreed with by others

APPENDIX F (Con't.)

OVERALL RATING

How would you rate the overall quality of this individual's performance in this exercise?

9 8 7 6 5 4 3 2 (1)

Excellent Average Poor

FOOTNOTES

FOOTNOTES

[1]It should be noted here that the discussion has been limited to H and L ratings preceding the A quality participant because it is in these sequences in which contrast effects have been found to have their greatest effects on ratings (Wexley, et al., 1972 and 1973).

[2]General correspondence between the fifteen subscales of the rating form, the suggested behaviors supplied to the participants, and the most prominent behaviors exhibited by the participants, was evidenced by the fact that for only a few of the participants was it necessary for a rater to exclude using a particular subscale.

BIBLIOGRAPHY

BIBLIOGRAPHY

Ansbacher, H. L. The history of the leaderless group discussion technique. Psychological Bulletin, 1951, 48, 383-390.

Bass, B. M. The leaderless group discussion. Psychological Bulletin, 1954, 51, 465-492.

Boehm, V. R. Negro-white differences in validity of employment and training selection procedures: Summary of research evidence. Journal of Applied Psychology, 1972, 56(1), 370-374.

Bray, D. W., & Grant, D. L. The assessment center in the measurement of potential for business management. Psychological Monographs, 1966, 80, (17, Whole No. 625).

Byham, W. C. The uses of assessment centers. (A summary of material presented at the Spring, 1969 meeting of the Executive Study Conference and material compiled in preparation for the conference.)

Cox, J. A., & Krumboltz, J. D. Racial bias in peer ratings of basic airmen. Sociometry, 1958, 21, 292-299.

De Jung, J. E., & Kaplan, H. Some differential effects of race of rater and ratee on early peer ratings of combat aptitude. Journal of Applied Psychology, 1962, 46(5), 370-374.

Greenwood, J. M., & McNamara, W. J. Interrater reliability in situational tests. Journal of Applied Psychology, 1967, 51(2), 101-106.

Guirintano, L. Stimulus generalization in humans and pigeons as a function of training and test stimulus parameters. Unpublished doctoral dissertation, University of Mississippi, 1973.

Hakel, M. D., Ohnesorge, J. P., & Dunnette, M. D. Interviewer evaluations of job applicants' resumes as a function of the qualifications of the immediately preceding applicants: An examination of contrast effects. Journal of Applied Psychology, 1970, 54(1), 27-30.

Helson, H. Adaptation-level as frame of reference for prediction of psychophysical data. American Journal of Psychology, 1947, 60(1), 1-29.

Helson, H. Adaptation-level Theory. New York: Harper & Row, 1964.

Helson, H., & Avant, L. L. Stimulus generalization as a function of contextual stimuli. Journal of Experimental Psychology, 1967, 73, 565-567.

Holmes, D. S., & Berkowitz, L. Some contrast effects in social perception. Journal of Abnormal and Social Psychology, 1961, 62(1), 150-152.

Hovland, C., Harvey, O., & Sherif, M. Assimilation and contrast effects in reaction to communication and attitude change. Journal of Abnormal and Social Psychology, 1957, 55, 244-252.

Jaffee, C. L. Effective Management Selection: The Analysis of Behavior by Simulation Techniques. Reading, Massachusetts: Addison-Wesley, 1971.

Levine, S. The effect of the induction of involvement on psychophysical judgments. Perception and Psychophysics, 1972, 11(1A).

Parrish, M., & Smith, K. Simultaneous brightness contrast as a function of perceptual set. Psychonomic Science, 1967, 7, 155-156.

Rowe, P. M. Order effects in assessment decisions. Journal of Applied Psychology, 1967, 51(2), 170-173.

Wexley, K. W., Sanders, R. E., & Yukl, G. A. Training interviewers to eliminate contrast effects in employment interviews. Journal of Applied Psychology, 1973, 57, 233-236.

Wexley, K. W., Yukl, G. A., Kovacs, S. Z., & Sanders, R. E. Importance of contrast effects in employment interviews. Journal of Applied Psychology, 1972, 56, 45-48.

Williams, M. Y., Ross, J., & Di Lollo, V. Sensory contrast effects in the judgment of lifted weights. Psychonomic Science, 1966, 6, 137-138.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.